RESPONSE

I. Status of the Claims

Claims 1-8 are pending and rejected under 35 U.S.C. § 101. As per revised 37 CFR 1.121 the claims and their current status are listed in **Exhibit A**.

II. Rejection of Claims 1-8 Under 35 U.S.C. § 101

The Action first rejects claims 1-8 under 35 U.S.C. § 101 because the claimed invention is allegedly not supported by either a credible, specific and substantial asserted utility or a well-established utility.

In Applicants' response to the First Office Action dated 12/06/01 (Paper no. 10) Applicants identified the fact that the sequences of the present invention demonstrated high levels of homology over shared domains with annotated sequences present in the leading scientific repository for biological sequence data (GenBank). For example the sequences described by Ishibashi *et al.*, 2000, AB013103, which was anotated as "Homo sapiens mRNA for MS4A5, complete cds") and Hulett *et al.*, 2001. Applicants respectfully submit further evidence that the amino acid sequence of SEQ ID NO:2 of the present invention is identical to a sequence. Accession No. Q9H3V2 (information provided as **Exhibit B**) has been annotated by third party scientists *wholly unaffiliated with Applicants* as encoding "Membrane-spanning 4-domains subfamily A member 5 (testis-expressed transmembrane 4 protein) (CD20 antigen -like 2)". Applicants assertion that the sequences of the present invention encode the CD20 antigen-like molecule Membrane-spanning 4-domains subfamily A member 5 (testis-expressed transmembrane 4 protein) (Q9H3V2) is supported by the evidence provided in **Exhibit C**, which contains an amino acid sequence comparison between SEQ ID NO:2 and the amino acid sequence of Q9H3V2. From this comparison, it can be seen that SEQ ID NO:2 is identical to the amino acid sequence of Q9H3V2. As stated in the specification (on pages 2, line 24) these proteins play a role as mediators of signal transduction.

The expression of the sequences of the present invention were described in the specification (page 3, line 5) as human testis cells and the activity of this protein is described in several publications

(Abstracts provided as **Exhibit D**). One is entitled "Identification of a new multigene four-transmembrane family (MS4A) related to CD20, Htm4 and beta subunit of the high-afinity IgE receptor" by Ishibashi ,*et al.*, 2001 (Gene *264*:87-93, 2001). Another is entitled "Isolation, tissue distribution, and chromosomal localization of a novel testis-specific human four-transmembrane gene related to CD20 and Fcepsilon RI-beta" by Hulett *et al.*, 2001 (Biochem Biophys Res Commun 280:374-9, 2001). A third and fourth are entitled " Identification of a CD20-, FcepsilonRIbeta-, and Htm4-related gene family: sixteen new MS4A family members expressed in human and mouse" by Liang and Tedder, 2001 (Genomics 72:119-27, 2001) and "Structural organization of the human MS4A gene cluster on Chromosome 11q12" by Liang, *et al.*, 2001 (Immunogenetics *53:*357-68, 2001).

These publications constitute evidence that clearly demonstrates that the proteins of the present invention have function and utility that are both accepted by those skilled in the art. As the legal test for utility simply involves an assessment of whether those skilled in the art would find any of the utilities described for the invention to be credible or believable. Clearly those of skill in the art would recognize that molecules that share identical amino acid sequences and tissue expression patterns would share protein structure and would thus also have the same function. This constitutes evidence that clearly supports the specifications assertion that sequences of the present invention have the function of a known protein.

Applicants submit that the presently claimed molecules have been shown to encode, as asserted in the specification as filed, a human CD20 antigen-like molecule, membrane-spanning 4-domains subfamily A member 5 (testis-expressed transmembrane 4 protein: MS4A5) which is known to the art. Thus, the present situation directly tracks Example 10 of the Revised Interim Utility Guidelines Training Materials (pages 53-55), which establishes that a rejection under 35 U.S.C. § 101 as allegedly lacking a patentable utility and under 35 U.S.C. § 112, first paragraph as allegedly unusable by the skilled artisan due to the alleged lack of patentable utility, is not proper when a full length sequence (such as the presently claimed sequence), and has a similarity score greater than 95% to a protein having a known function (such as the 100% identity between the presently claimed sequences and those of the cited protein (Q9H3V2). Further evidence of utility of the presently claimed polynucleotide, although only one is needed to meet the requirements of 35 U.S.C. § 101 (*Raytheon v. Roper*, 220 USPQ 592

(Fed. Cir. 1983); *In re Gottlieb*, 140 USPQ 665 (CCPA 1964); *In re Malachowski*, 189 USPQ 432 (CCPA 1976); *Hoffman v. Klaus*, 9 USPQ2d 1657 (Bd. Pat. App. & Inter. 1988)), is the specific utility the present nucleotide sequence has in determining the genomic structure of the corresponding human chromosome, for example mapping the protein encoding regions as described in the specification. As evidence supporting Applicants assertions of the specific utility of the sequences of the present invention in localizing the specific region of the human chromosome and identification of functionally active intron/exon splice junctions is the information provided as **Exhibit E**. This is the result of overlaying the sequence of SEQ ID NO:1 of the present invention and the identified human genomic sequence. By doing this, one is able to identify the portions of the genome that encode the present invention. If these regions of the genome are non-contiguous, this is indicative of individual exons. The results of such an analysis indicates that the sequence of the present invention is encoded by 5 exons spread non-contiguously along a region of human chromosome 11, which is contained within Accession no. AP001034.5 in the BAC clone RP11-729B4. Thus clearly one would not simply be able to identify the 5 protein encoding exons that make up the sequence of the present intention from within the large genomic sequence. Nor, would one be able to map the protein encoding regions identified specifically by the sequences of the present invention without knowing exactly what those specific sequences were. Additionally, it should be noted that the human MS4A5 gene also maps to the same region of human chromosome 11 (at approximately 11q12). This further supports Applicant's position that the sequences of the present invention encodes human MS4A5

Therefore, clearly, the present polynucleotide provides exquisite specificity in localizing the specific region of the human chromosome containing the gene encoding the given polynucleotide, a utility not shared by virtually any other nucleic acid sequence. In fact, it is this specificity that makes this particular sequence so useful. Early gene mapping techniques relied on methods such as Giemsa staining to identify regions of chromosomes. However, such techniques produced genetic maps with a resolution of only 5 to 10 megabases, far too low to be of much help in identifying specific genes involved in disease. The skilled artisan readily appreciates the significant benefit afforded by markers that map a specific locus of the human genome, such as the present nucleic acid sequence.

Only a minor percentage of the genome actually encodes exons, which in turn encode amino acid sequences. The presently claimed polynucleotide sequence provides biologically validated

empirical data (*e.g.*, showing which sequences are transcribed, spliced, and polyadenylated) that *specifically* defines that portion of the corresponding genomic locus that actually encodes exon sequence. Equally significant is that the claimed polynucleotide sequence defines how the encoded exons are actually spliced together to produce an active transcript (*i.e.*, the described sequences are useful for functionally defining exon splice-junctions). The Applicants respectfully submit that the practical scientific value of expressed, spliced, and polyadenylated mRNA sequences is readily apparent to those skilled in the relevant biological and biochemical arts. The presently claimed polynucleotide sequence defines a biologically validated sequence that provides a unique and specific resource for mapping the genome.

An additional utility includes the use of the presently claimed polynucleotides on DNA chips. Further, the Action seems to be requiring Applicants to identify the biological role of the nucleic acid or function of the protein encoded by the presently claimed polynucleotides before the present sequences can be used in gene chip applications that meet the requirements of § 101. Applicants respectfully point out that knowledge of the exact function or role of the presently claimed sequence is not required to track expression patterns using a DNA chip. Given the widespread utility of such "gene chip" methods using *public domain* gene sequence information, there can be little doubt that the use of the presently described *novel* sequences would have great utility in such DNA chip applications. The claimed sequence provides a specific marker of the human genome (see evidence below), and that such specific markers are targets for discovering drugs that are associated with human disease. Thus, those skilled in the art would instantly recognize that the present nucleotide sequence would be an ideal, novel candidate for assessing gene expression using, for example, DNA chips, as the specification details. Such "DNA chips" clearly have utility, as evidenced by hundreds of issued U.S. Patents, as exemplified by U.S. Patent Nos. 5,445,934, 5,556,752, 5,744,305, as well as more recently issued U.S. Patent Nos. 5,837,832, 6,156,501 and 6,261,776. Accordingly, the present sequence has a specific utility in such DNA chip applications. Clearly, compositions that enhance the utility of such DNA chips, such as the presently claimed nucleotide sequence, must also be useful.

Additionally, since only a small percentage of the genome (2-4%) actually encodes exons, which in-turn encode amino acid sequences. Thus, not all human genomic DNA sequences are useful in such gene chip applications. Thus, the present claims clearly meet the requirements of 35 U.S.C. §

101. It has been clearly established that a statement of utility in a specification must be accepted absent reasons why one skilled in the art would have reason to doubt the objective truth of such statement. *In re Langer*, 503 F.2d 1380, 1391, 183 USPQ 288, 297 (CCPA, 1974); *In re Marzocchi*, 439 F.2d 220, 224, 169 USPQ 367, 370 (CCPA, 1971).

Evidence of the "real world" substantial utility of the present invention is further provided by the fact that there is an entire industry established based on the use of gene sequences or fragments thereof in a gene chip format. Perhaps the most notable gene chip company is Affymetrix. However, there are many companies which have, at one time or another, concentrated on the use of gene sequences or fragments, in gene chip and non-gene chip formats, for example: Gene Logic, ABI-Perkin-Elmer, HySeq and Incyte. In addition, one such company, Rosetta Inpharmatics, was viewed to have such "real world" value that it was acquired by large pharmaceutical company, Merck & Co., for substantial sums of money (net equity value of the transaction was $620 million). The "real world" substantial industrial utility of gene sequences or fragments would, therefore, appear to be widespread and well established. Clearly, persons of skill in the art, as well as venture capitalists and investors, readily recognize the utility, both scientific and commercial, of genomic data in general, and specifically human genomic data. Billions of dollars have been invested in the human genome project, resulting in useful genomic data (see, *e.g.*, Venter *et al.*, 2001, Science *291*:1304). The results have been a stunning success as the utility of human genomic data has been widely recognized as a great gift to humanity (see, *e.g.*, Jasny and Kennedy, 2001, Science *291*:1153). Clearly, the usefulness of human genomic data, such as the presently claimed nucleic acid molecules, is substantial and credible (worthy of billions of dollars and the creation of numerous companies focused on such information) and well-established (the utility of human genomic information has been clearly understood for many years). The sequences of the present invention have particularly specific utility in DNA gene chip based analysis as they have been identified to contain several coding region nucleotide polymorphisms (see above), thus increasing their utility in DNA gene chip based analysis.

Finally, the Examiner is requested to consider the issue of due process. Applicants understanding is that issued United States patents retain a legal presumption of validity which in this case indicates that the inventions claimed in the cited patents are *legally presumed* to be in full compliance with the provisions of 35 U.S.C. sections 101, 102, 103, and 112. Applicants respectfully submit that,

absent a change in the law as enacted by Congress and signed by the President, it is improper for the Examiner to hold Applicants' invention to a different legal standard of patentability. Given the rapid pace of development in the biotechnology arts, it is difficult for the Applicants to understand how an invention fully disclosed and free of prior art at the time the present application was filed, could somehow retain *less* utility and be *less* enabled than inventions in the cited issued U.S. patents (which were filed during a time when the level of skill in the art was clearly lower). Simply put, Applicants invention is *more* enabled and retains *at least as much* utility as the inventions described in the claims of the U.S. patents of record. Any argument to the contrary is at best arbitrary and at worst capricious. Absent authority provided by an act of Congress or Executive order, arbitrary or capricious conduct by an administrative office the U.S. government has historically proven to conflict with the provisions of the U.S. Constitution. The Patent Office does not have the authority to rewrite U.S. law. However, the Patent Office does have a Constitutional obligation to administer U.S. law in an unbiased and procedurally consistent manner. That is what the Applicants are respectfully requesting the Examiner to consider in the present matter. As the issued U.S. Patents cited above are presumed to meet all of the requirements for patentability, including 35 U.S.C. §§ 101 and 112, first paragraph, Applicants respectfully submit that the presently claimed polynucleotide must also meet the requirements of 35 U.S.C. § 101.

Thus in summary, Applicants submit that the presently claimed molecules have been shown to encode, as asserted in the specification as filed, MS4A5, whose biological function is known to the art. Thus, the present situation directly tracks Example 10 of the Revised Interim Utility Guidelines Training Materials (pages 53-55), which establishes that a rejection under 35 U.S.C. § 101 as allegedly lacking a patentable utility and under 35 U.S.C. § 112, first paragraph as allegedly unusable by the skilled artisan due to the alleged lack of patentable utility, is not proper when a full length sequence (such as the presently claimed sequence), and has a similarity score greater than 95% to a protein having a known function (such as the 100% identity between the presently claimed sequences and those of the cited protein). Furthermore this response has described a series of additional substantial, specific, credible and well-established utilities for the present invention. Therefore, Applicants submit that as the presently claimed sequence molecules have been shown to have a substantial, specific, credible and well-established utility, the rejection of the claims under 35 U.S.C. § 101 has been overcome. Thus,

Applicants respectfully request that the rejection be withdrawn.

V. Rejection of Claims 1-8 Under 35 U.S.C. § 112, First Paragraph

The Action rejects claims 1-8 under 35 U.S.C. § 112, first paragraph, as allegedly the claimed invention is not supported by either a specific or substantial asserted utility or a well established utility allegedly one skilled in the art clearly would not know how to use the claimed invention. Applicants respectfully submit that claims 1-8 have been shown to have "a specific, substantial, and credible utility", as detailed in the section IV above. Therefore, one skilled in the art would clearly know how to use the claimed invention and Applicants therefore request that the rejection of claims 1-8 under 35 U.S.C. § 112, first paragraph, be withdrawn.

VI. Conclusion

The present document is a full and complete response to the Action. In conclusion, Applicants submit that, in light of the foregoing remarks, the present case is in condition for allowance, and such favorable action is respectfully requested. Should Examiner Li have any questions or comments, or believe that certain amendments of the claims might serve to improve their clarity, a telephone call to the undersigned Applicants' representative is earnestly solicited.

Respectfully submitted,

August 1, 2003
Date

Lance K. Ishimoto Reg. No. No. 41,866
Agent for Applicants

LEXICON GENETICS INCORPORATED
(281) 863-3333

24231
PATENT TRADEMARK OFFICE

Exhibit A

Status of Claims in

U.S. Patent Application Ser. No. 09/735,712

1.(previously presented) An isolated nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1.

2.(previously presented) An isolated nucleic acid molecule comprising a sequence that:

(a) encodes the amino acid sequence shown in SEQ ID NO: 2; and

(b) hybridizes under highly stringent conditions with wash conditions of 0.1xSSC/0.1%SDS at 68°C to the nucleotide sequence of SEQ ID NO: 1 or the complement thereof.

3.(original) An isolated nucleic acid molecule comprising a nucleotide sequence that encodes the amino acid sequence shown in SEQ ID NO: 2.

4.(original) An isolated nucleic acid molecule comprising a nucleotide sequence that encodes the amino acid sequence shown in SEQ ID NO: 8.

5.(previously presented) A recombinant expression vector comprising the isolated nucleic acid molecule of claim 3.

6.(previously presented) A host cell comprising the recombinant expression vector of claim 5.

7.(previously presented) A recombinant expression vector comprising the isolated nucleic acid molecule of claim 4.

8.(previously presented) A host cell comprising the recombinant expression vector of claim 7.

NCBI Entrez Protein

PubMed Nucleotide Protein Genome Structure PMC Taxonomy OMIM Boo

Search Protein for Go Clear

Limits Preview/Index History Clipboard Details

Display default Show: 20 Send to File Get Subsequence

☐ 1: Q9H3V2. Membrane-spanning...[gi:29611830] BLink, Domains, Links

```
LOCUS       Q9H3V2                  200 aa            linear   PRI 15-SEP-2003
DEFINITION  Membrane-spanning 4-domains subfamily A member 5 (Testis-expressed
            transmembrane 4 protein) (CD20 antigen-like 2).
ACCESSION   Q9H3V2
VERSION     Q9H3V2  GI:29611830
DBSOURCE    swissprot: locus M4A5_HUMAN, accession Q9H3V2;
            class: standard.
            extra accessions:Q9BZH1,created: Sep 15, 2003.
            sequence updated: Sep 15, 2003.
            annotation updated: Sep 15, 2003.
            xrefs: gi: 11559213, gi: 11559214, gi: 13649400, gi: 13649401, gi:
            12698681, gi: 12698682, gi: 20988638, gi: 20988639, gi: 25392180
            xrefs (non-sequence databases): GenewHGNC:13374, MIM 606499,
            InterProIPR007237, PfamPF04103
KEYWORDS    Receptor; Transmembrane; Multigene family.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (residues 1 to 200)
  AUTHORS   Ishibashi,K., Suzuki,M., Sasaki,S. and Imai,M.
  TITLE     Identification of a new multigene four-transmembrane family (MS4A)
            related to CD20, HTm4 and beta subunit of the high-affinity IgE
            receptor
  JOURNAL   Gene 264 (1), 87-93 (2001)
  MEDLINE   21142397
  REMARK    SEQUENCE FROM N.A.
REFERENCE   2  (residues 1 to 200)
  AUTHORS   Liang,Y. and Tedder,T.F.
  TITLE     Identification of a CD20-, FcepsilonRIbeta-, and HTm4-related gene
            family: sixteen new MS4A family members expressed in human and
            mouse
  JOURNAL   Genomics 72 (2), 119-127 (2001)
  MEDLINE   21295030
  REMARK    SEQUENCE FROM N.A.
REFERENCE   3  (residues 1 to 200)
  AUTHORS   Hulett,M.D., Pagler,E., Hornby,J.R., Hogarth,P.M., Eyre,H.J.,
            Baker,E., Crawford,J., Sutherland,G.R., Ohms,S.J. and Parish,C.R.
  TITLE     Isolation, tissue distribution, and chromosomal localization of a
            novel testis-specific human four-transmembrane gene related to CD20
            and FcepsilonRI-beta
  JOURNAL   Biochem. Biophys. Res. Commun. 280 (1), 374-379 (2001)
  MEDLINE   21092614
  REMARK    SEQUENCE FROM N.A.
REFERENCE   4  (residues 1 to 200)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
```

Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S.N., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J.M. and Marra,M.A.

TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

MEDLINE 22388257

REMARK SEQUENCE FROM N.A.
TISSUE=Brain

COMMENT ----------
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute.
The original entry is available from http://www.expasy.ch/sprot and http://www.ebi.ac.uk/sprot
----------.
[FUNCTION] May be involved in signal transduction as a component of a multimeric receptor complex.
[SUBCELLULAR LOCATION] Integral membrane protein.
[TISSUE SPECIFICITY] Expressed at high level in the testis. Detected also in the pancreas, heart and in the brain.
[SIMILARITY] Belongs to the MS4A family.

FEATURES Location/Qualifiers

source 1..200
/organism="Homo sapiens"
/db_xref="taxon:9606"

gene 1..200
/gene="MS4A5"
/note="synonyms: TETM4, CD20L2"

Protein 1..200
/gene="MS4A5"
/product="Membrane-spanning 4-domains subfamily A member 5"

Region 1..52
/gene="MS4A5"
/region_name="Domain"
/note="CYTOPLASMIC (POTENTIAL)."

Region 53..73
/gene="MS4A5"
/region_name="Transmembrane region"
/note="POTENTIAL."

Region 74..80
/gene="MS4A5"
/region_name="Domain"
/note="CYTOPLASMIC (POTENTIAL)."

Region 81..101
/gene="MS4A5"

```
                     /region_name="Transmembrane region"
                     /note="POTENTIAL."
     Region          102..120
                     /gene="MS4A5"
                     /region_name="Domain"
                     /note="CYTOPLASMIC (POTENTIAL)."
     Region          121..141
                     /gene="MS4A5"
                     /region_name="Transmembrane region"
                     /note="POTENTIAL."
     Region          142..159
                     /gene="MS4A5"
                     /region_name="Domain"
                     /note="CYTOPLASMIC (POTENTIAL)."
     Region          160..180
                     /gene="MS4A5"
                     /region_name="Transmembrane region"
                     /note="POTENTIAL."
     Region          163
                     /gene="MS4A5"
                     /region_name="Conflict"
                     /note="F -> L (IN REF. 3)."
     Region          181..200
                     /gene="MS4A5"
                     /region_name="Domain"
                     /note="CYTOPLASMIC (POTENTIAL)."
ORIGIN
        1 mdsstahspv flvfppeita seyestelsa ttfstqsplq klfarkmkil gtiqilfgim
       61 tfsfgviflf tllkpyprfp fiflsgypfw gsvlfinsga fliavkrktt etliilsrim
      121 nflsalgaia giilltfgfi ldqnyicgys hqnsqckavt vlflgilitl mtfsiielfi
      181 slpfsilgch sedcdceqcc
//
```

```
 FASTA searches a protein or DNA sequence data bank
 version 3.3t05 March 30, 2000
Please cite:
 W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

/tmp/fastaCAA1JaGhY: 201 aa
 >LEX 109 SEQ ID NO:2
 vs  /tmp/fastaDAA2JaGhY library
searching /tmp/fastaDAA2JaGhY library

    200 residues in     1 sequences

FASTA (3.34 January 2000) function [optimized, BL50 matrix (15:-5)] ktup: 2
 join: 36, opt: 24, gap-pen: -12/ -2, width:  16
 Scan time:  0.017
The best scores are:                                                opt
gi|29611830|sp|Q9H3V2|M4A5_HUMAN Membrane-spannin  ( 200) 1285

>>gi|29611830|sp|Q9H3V2|M4A5_HUMAN Membrane-spanning 4-d  (200 aa)
 initn: 1285 init1: 1285 opt: 1285
Smith-Waterman score: 1285;  98.500% identity in 200 aa overlap (1-200:1-200)

               10        20        30        40        50        60
LEX     MDSSTAHSPVFLVFPPEITASEYESTELSATTFSTQSPLQKLFARKMKILGTIQILFGIM
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
gi|296 MDSSTAHSPVFLVFPPEITASEYESTELSATTFSTQSPLQKLFARKMKILGTIQILFGIM
               10        20        30        40        50        60

               70        80        90       100       110       120
LEX     TFSFGVIFLFXLLKPYPRFPFIFLSGYPFWGSVLFINSGAFLIAVKRKTTETLIILSRIM
        ::::::::::.::::::::::::::::::::::::::::::::::::::::::::::::::
gi|296 TFSFGVIFLFTLLKPYPRFPFIFLSGYPFWGSVLFINSGAFLIAVKRKTTETLIILSRIM
               70        80        90       100       110       120

              130       140       150       160       170       180
LEX     NXLSALXAIAGIILLTFGFILDQNYICGYSHQNSQCKAVTVLFLGILITLMTFSIIELFI
        : :::: ::::::::::::::::::::::::::::::::::::::::::::::::::::::
gi|296 NFLSALGAIAGIILLTFGFILDQNYICGYSHQNSQCKAVTVLFLGILITLMTFSIIELFI
              130       140       150       160       170       180

              190       200
LEX     SLPFSILGCHSEDCDCEQCC*
        ::::::::::::::::::::
gi|296 SLPFSILGCHSEDCDCEQCC
              190       200


201 residues in 1 query   sequences
200 residues in 1 library sequences
 Scomplib [version 3.3t05 March 30, 2000]
 start: Tue Jul 29 16:52:47 2003 done: Tue Jul 29 16:52:47 2003
 Scan time:  0.017 Display time:  0.050

Function used was FASTA
```

EXHIBIT "D"

LEX 109 Abstracts

Gene. 2001 Feb 7;264(1):87-93. Related Articles, Links

ELSEVIER SCIENCE
FULL-TEXT ARTICLE

Identification of a new multigene four-transmembrane family (MS4A) related to CD20, HTm4 and beta subunit of the high-affinity IgE receptor.

Ishibashi K, Suzuki M, Sasaki S, Imai M.

Department of Pharmacology, Jichi Medical School, Minamikawachi, Kawachi, 329-0498, Tochigi, Japan. kishiba@jichi.ac.jp

We report here the cloning of eight new cDNAs that encode a family of proteins related to the B-cell-specific antigen CD20, a hematopoietic-cell-specific protein HTm4, and high affinity IgE receptor beta chain (FcvarepsilonRIbeta). They include four clones from human, and another four clones from mouse. They share similar structure (four transmembrane domains) with amino acid identities of 25-40%. Therefore, they represent distinct genes and comprise a gene superfamily. This superfamily is now named membrane-spanning four-domains, subfamily A (the approved symbol is MS4A) to distinguish them from tetraspanins with similar structure. The highest homologies among these proteins are found in the transmembrane domains, especially in the first and second transmembrane domains, and conserved residues are also recognized in the inter-transmembrane domains. In northern blot, they were mostly expressed in lymphoid tissues: thymus and spleen. However, some were expressed in nonlymphoid tissues including brain, heart, kidney, liver, testis, lung, GI tracts, and pancreas. They may represent proteins functioning either directly as ligand-gated ion channels or as essential components of such channels. The identification of this relatively large gene family in various tissues will allow the further elucidation of physiological significance of this gene family, that is currently unclear.

PMID: 11245982 [PubMed - indexed for MEDLINE]

Biochem Biophys Res Commun. 2001 Jan 12;280(1):374-9.

Isolation, tissue distribution, and chromosomal localization of a novel testis-specific human four-transmembrane gene related to CD20 and FcepsilonRI-beta.

Hulett MD, Pagler E, Hornby JR, Hogarth PM, Eyre HJ, Baker E, Crawford J, Sutherland GR, Ohms SJ, Parish CR.

Division of Immunology and Cell Biology, John Curtin School of Medical Research, ANU, P.O. Box 344, Canberra, ACT 2601, Australia.
mark.hulett@anu.edu.au

CD20 and the beta subunit of the high affinity receptor for IgE (FcepsilonRIbeta) are related four-transmembrane molecules that are expressed on the surface of hematopoietic cells and play crucial roles in signal transduction. Herein, we report the identification and characterization of a human gene, TETM4, that encodes a novel four-transmembrane protein related to CD20 and FcepsilonRIbeta. The predicted TETM4 protein is 200 amino acids and contains four putative transmembrane regions, N- and C-terminal cytoplasmic domains, and three inter-transmembrane loop regions. TETM4 shows 31.0 and 23.2% overall identity with CD20 and FcepsilonRIbeta respectively, with the highest identity in the transmembrane regions, whereas the N- and C-termini and inter-transmembrane loops are more divergent. Northern blot and RT-PCR analysis suggest that TETM4 mRNA has a highly restricted tissue distribution, being expressed selectively in the testis. Using fluorescence in situ hybridization and radiation hybrid analysis, the TETM4 gene has been localized to chromosome 11q12. The genes for CD20 and FcepsilonRIbeta have also been mapped to the same region of chromosome 11 (11q12-13.1), suggesting that these genes have evolved by duplication to form a family of four-transmembrane genes. TETM4 is the first nonhematopoietic member of the CD20/FcepsilonRIbeta family, and like its hematopoietic-specific relatives, it may be involved in signal transduction as a component of a multimeric receptor complex.

PMID: 11162526 [PubMed - indexed for MEDLINE]

Genomics. 2001 Mar 1;72(2):119-27.

Identification of a CD20-, FcepsilonRIbeta-, and HTm4-related gene family: sixteen new MS4A family members expressed in human and mouse.

Liang Y, Tedder TF.

Department of Immunology, Duke University Medical Center, Durham, North Carolina 27710, USA.

CD20, high-affinity IgE receptor beta chain (FcepsilonRIbeta), and HTm4 are structurally related cell-surface proteins expressed by hematopoietic cells. In the current study, 16 novel human and mouse genes that encode new members of this nascent protein family were identified. All family members had at least four potential membrane-spanning domains, with N- and C-terminal cytoplasmic domains. This family was therefore named the membrane-spanning 4A gene family, with at least 12 subgroups (MS4A1 through MS4A12) currently representing at least 21 distinct human and mouse proteins. Each family member had unique patterns of expression among hematopoietic cells and nonlymphoid tissues. Four of the 6 human MS4A genes identified in this study mapped to chromosome 11q12-q13.1 along with CD20, FcepsilonRIbeta, and HTm4. Thus, like CD20 and FcepsilonRIbeta, the other MS4A family members are likely to be components of oligomeric cell surface complexes that serve diverse signal transduction functions.

PMID: 11401424 [PubMed - indexed for MEDLINE]

Immunogenetics. 2001 Jul;53(5):357-68.

Structural organization of the human MS4A gene cluster on Chromosome 11q12.

Liang Y, Buckley TR, Tu L, Langdon SD, Tedder TF.

Department of Immunology, Room 353 Jones Building, Research Drive, Duke University Medical Center, Durham, NC 27710, USA.

CD20, the high-affinity IgE receptor beta chain (FcepsilonRIbeta), and HTm4 are structurally related cell surface proteins expressed by hematopoietic cells. Recently, 16 novel human and mouse genes were identified that encode new members of this nascent protein family that we have named the membrane-spanning 4A gene family, with at least 12 subgroups (MS4A1-MS4A12). In the current study, we identified three additional human MS4A genes: MS4A4E, MS4A6E, and MS4A10. All family members have at least four potential transmembrane domains and N- and C-terminal cytoplasmic domains encoded by distinct exons, except MS4A6E which contains two transmembrane domains. Otherwise, the 12 currently identified MS4A genes share common structural features and similar intron/exon splice boundaries, and are clustered along an approximately 600-kb region of Chromosome 11q12. In contrast to other MS4A genes, MS4A4E, MS4A6E, and MS4A10 transcripts were rare and not detected among hematopoietic cells and most nonlymphoid tissues. Sequence polymorphisms were identified in the MS4A6E gene and common splice variants were observed for the MS4A4A, MS4A5, MS4A6A, and MS4A7 genes. Thus, the MS4A family currently includes 24 distinct human and mouse genes. Like CD20 and FcepsilonRIbeta, the 10 other human MS4A family members are likely to be components of oligomeric cell surface complexes involved in signal transduction in diverse cell lineages.

PMID: 11486273 [PubMed - indexed for MEDLINE]

Home **Paracel BLAST Results** Help

MEGABLAST 1.2.3-Paracel [2001-11-20]

Reference:
Zheng Zhang, Scott Schwartz, Lukas Wagner, and Webb Miller (2000),
"A greedy algorithm for aligning DNA sequences",
J Comput Biol 2000; 7(1-2):203-14.
Database: Homo_sapiens.latestgp.fa
26,679 sequences; 200,800,637,119 total letters

Query= Orf1
(603 letters)

Sequences producing significant alignments:	Score (bits)	E Value
AP001034.5.1.167934	304	3e-80
AP003127.2.1.161238	304	3e-80

>AP001034.5.1.167934
Length = 167934

Score = 304 bits (153), Expect = 3e-80
Identities = 153/153 (100%)
Strand = Plus / Plus

```
Query: 1      atggattcaagcaccgcacacagtccggtgtttctggtatttcctccagaaatcactgct 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 22721 atggattcaagcaccgcacacagtccggtgtttctggtatttcctccagaaatcactgct 22780

Query: 61     tcagaatatgagtccacagaactttcagccacgacctttcaactcaaagccccttgcaa 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 22781 tcagaatatgagtccacagaactttcagccacgacctttcaactcaaagccccttgcaa 22840

Query: 121    aaattatttgctagaaaaatgaaaatcttaggg 153
              |||||||||||||||||||||||||||||||
Sbjct: 22841 aaattatttgctagaaaaatgaaaatcttaggg 22873
```

Score = 300 bits (151), Expect = 5e-79
Identities = 153/155 (98%)
Strand = Plus / Plus

```
Query: 339    gataatattgagccgaataatgaatyttcttagtgccctgrgagcaatagctggaatcat 398
              |||||||||||||||||||||||| ||||||||||||| ||||||||||||||||||
Sbjct: 26810 gataatattgagccgaataatgaattttcttagtgccctgggagcaatagctggaatcat 26869

Query: 399    tctcctcacatttggtttcatcctagatcaaaactacatttgtggttattctcaccaaaa 458
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 26870 tctcctcacatttggtttcatcctagatcaaaactacatttgtggttattctcaccaaaa 26929
```

```
Query: 459   tagtcagtgtaaggctgttactgtcctgttcttgg 493
             |||||||||||||||||||||||||||||||||
Sbjct: 26930 tagtcagtgtaaggctgttactgtcctgttcttgg 26964
```

Score = 255 bits (128), Expect = 3e-65
Identities = 129/130 (99%)
Strand = Plus / Plus

```
Query: 153   gactatccagatcctgtttggaattatgacctttttcttttggagttatcttccttttcac 212
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 23841 gactatccagatcctgtttggaattatgacctttttcttttggagttatcttccttttcac 23900
```

(2)

```
Query: 213   yttgttaaaaccatatccaaggtttccctttatatttctttcaggatatccattctgggg 272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 23901 cttgttaaaaccatatccaaggtttccctttatatttctttcaggatatccattctgggg 23960
```

```
Query: 273   ctctgttttg 282
              ||||||||||
Sbjct: 23961 ctctgttttg 23970
```

Score = 223 bits (112), Expect = 1e-55
Identities = 112/112 (100%)
Strand = Plus / Plus

```
Query: 492   gggaattttgattacattgatgactttcagcattattgaattattcatttctctgccttt 551
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 40689 gggaattttgattacattgatgactttcagcattattgaattattcatttctctgccttt 40748
```

(5)

```
Query: 552   ctcaattttggggtgccactcagaggattgtgattgtgaacaatgttgttga 603
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 40749 ctcaattttggggtgccactcagaggattgtgattgtgaacaatgttgttga 40800
```

Score = 115 bits (58), Expect = 3e-23
Identities = 58/58 (100%)
Strand = Plus / Plus

(3)

```
Query: 282   gttcattaattctggagccttcctaattgcagtgaaaagaaaaaccacagaaactctg 339
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 25456 gttcattaattctggagccttcctaattgcagtgaaaagaaaaaccacagaaactctg 25513
```

```
>AP003127.2.1.161238
          Length = 161238
```

Score = 304 bits (153), Expect = 3e-80
Identities = 153/153 (100%)

```
 Strand = Plus / Plus

Query: 1      atggattcaagcaccgcacacagtccggtgtttctggtatttcctccagaaatcactgct 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 133596 atggattcaagcaccgcacacagtccggtgtttctggtatttcctccagaaatcactgct 133655


Query: 61     tcagaatatgagtccacagaactttcagccacgacctttttcaactcaaagccccttgcaa 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 133656 tcagaatatgagtccacagaactttcagccacgacctttttcaactcaaagccccttgcaa 133715


Query: 121    aaattatttgctagaaaaatgaaaatcttaggg 153
               |||||||||||||||||||||||||||||||
Sbjct: 133716 aaattatttgctagaaaaatgaaaatcttaggg 133748


 Score =  300 bits (151), Expect = 5e-79
 Identities = 153/155 (98%)
 Strand = Plus / Plus


Query: 339    gataatattgagccgaataatgaatyttcttagtgccctgrgagcaatagctggaatcat 398
               ||||||||||||||||||||||| ||||||||||||| ||||||||||||||||||
Sbjct: 137685 gataatattgagccgaataatgaattttcttagtgccctgggagcaatagctggaatcat 137744


Query: 399    tctcctcacatttggtttcatcctagatcaaaactacatttgtggttattctcaccaaaa 458
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 137745 tctcctcacatttggtttcatcctagatcaaaactacatttgtggttattctcaccaaaa 137804


Query: 459    tagtcagtgtaaggctgttactgtcctgttcttgg 493
               |||||||||||||||||||||||||||||||||
Sbjct: 137805 tagtcagtgtaaggctgttactgtcctgttcttgg 137839


 Score =  255 bits (128), Expect = 3e-65
 Identities = 129/130 (99%)
 Strand = Plus / Plus


Query: 153    gactatccagatcctgtttggaattatgacctttttcttttggagttatcttccttttcac 212
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 134716 gactatccagatcctgtttggaattatgacctttttcttttggagttatcttccttttcac 134775


Query: 213    yttgttaaaaccatatccaaggtttccctttatatttctttcaggatatccattctgggg 272
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 134776 cttgttaaaaccatatccaaggtttccctttatatttctttcaggatatccattctgggg 134835


Query: 273    ctctgttttg 282
               ||||||||||
Sbjct: 134836 ctctgttttg 134845
```

```
 Score =  223 bits (112), Expect = 1e-55
 Identities = 112/112 (100%)
 Strand = Plus / Plus

Query: 492    gggaattttgattacattgatgactttcagcattattgaattattcatttctctgccttt 551
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 151569 gggaattttgattacattgatgactttcagcattattgaattattcatttctctgccttt 151628


Query: 552    ctcaattttggggtgccactcagaggattgtgattgtgaacaatgttgttga 603
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 151629 ctcaattttggggtgccactcagaggattgtgattgtgaacaatgttgttga 151680


 Score =  115 bits (58), Expect = 3e-23
 Identities = 58/58 (100%)
 Strand = Plus / Plus


Query: 282    gttcattaattctggagccttcctaattgcagtgaaaagaaaaaccacagaaactctg 339
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct: 136331 gttcattaattctggagccttcctaattgcagtgaaaagaaaaaccacagaaactctg 136388


  Database: Homo_sapiens.latestgp.fa
    Posted date:  Jul 8, 2003 12:51 PM
  Number of letters in database: 200,800,637,119
  Number of sequences in database:  26,679

Lambda     K      H
    1.38    0.711     1.31

Gapped
Lambda     K      H
    1.38    0.711     1.31


Matrix: blastn matrix:1 -3
Gap Penalties: Existence: 0, Extension: 0
Number of Hits to DB: 0
length of query: 1208
length of database: 200,800,637,119
effective HSP length: 21
effective length of query: 582
effective search space used:        0
T: 0
A: 0
X1: 0 ( 0.0 bits)
X2: 20 (39.8 bits)
S1: 12 (24.4 bits)
S2: 37 (74.1 bits)
```